

Error estimates on a finite volume method for diffusion problems with interface on Eulerian grids

Jie Peng^a, Shi Shu^{a,b}, HaiYuan Yu^{a,*}, Chunsheng Feng^{a,c}, Mingxian Kan^d, Ganghua Wang^d

^a*School of Mathematics and Computational Science, Xiangtan University, Xiangtan 411105, China*

^b*Hunan Key Laboratory for Computation and Simulation in Science and Engineering, Xiangtan University, Xiangtan 411105, China*

^c*Guangdong Provincial Engineering Technology Research Center for Data Science, Guangzhou 510631, China*

^d*Institute of Fluid Physics, CAEP, P. O. Box 919 105, Mianyang 621900, China*

Abstract

The finite volume methods are frequently employed in the discretization of diffusion problems with interface. In this paper, we firstly present a vertex-centered MACH-like finite volume method for solving stationary diffusion problems with strong discontinuity and multiple material cells on the Eulerian grids. This method is motivated by Frese [No. AMRC-R-874, Mission Research Corp., Albuquerque, NM, 1987]. Then, the local truncation error and global error estimates of the degenerate five-point MACH-like scheme are derived by introducing some new techniques. Especially under some assumptions, we prove that this scheme can reach the asymptotic optimal error estimate $O(h^2 |\ln h|)$ in the maximum norm. Finally, numerical experiments verify theoretical results.

Keywords: diffusion problems with interface; finite volume method; Eulerian grids; error estimates
2010 MSC: 65N08, 65N12, 65N15

1. Introduction

Diffusion problems with interface are widely applied in multi-fluid hydrodynamic, fluid-solid coupling mechanics and many other scientific and engineering computation fields. The finite volume method (FVM), which presents local conservation and obvious advantage to handle physical models with complex characteristics very well, becomes an important discretization method for solving partial differential equations.

The finite volume methods (FVMs) based on Lagrangian and Eulerian grids are two commonly used methods for solving diffusion problems. The moving interface can be accurately described as we use the former, but the calculation is hard to execute on highly distorted grids. There are a lot of researches interested in these FVMs, e.g. [1, 2, 3, 4, 5, 6]. An advantage of the FVMs in the Eulerian frame is the reasonable shape of the computational grids, such as the uniform grids. But for the diffusion problems with strong discontinuity on the internal interface, the difficulty lies in dealing with the cells involving multiple material properties ([7]). Many researchers investigated this kind of FVMs, e.g. [8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19]. Ewing constructed an immersed finite volume element method (FVEM) on a uniform triangle grid, and presented the optimal error estimate in the

*Corresponding author

Email addresses: xtu_pengjie@163.com (Jie Peng), shushi@xtu.edu.cn (Shi Shu), spring@xtu.edu.cn (HaiYuan Yu), fengchsh@xtu.edu.cn (Chunsheng Feng), kanmx@caep.cn (Mingxian Kan), wanggh@caep.cn (Ganghua Wang)

energy norm ([8, 9]). Shu established the superconvergence theory of a bilinear FVM for diffusion problems with smooth coefficients on the rectangular grids ([10, 11]). Ma presented a recovery-based posterior error estimator for FVMs to solve elliptic interface problems in [12]. There also have a lot of works on the quadrilateral grids. For example, Oevermann focused on the two-dimensional (2D) and three-dimensional (3D) problems with discontinuous fluxes across the interface, and constructed a linear immersed FVM in [13] and [14]. Ewing constructed a series of FVMs with different average methods for diffusion problems under homogeneous jump conditions, and numerical experiments were carried on to confirm the approximation order of these methods in [15]. Luce derived a kind of FVM by using a decomposition technique in [16]. Wang constructed a fourth-order compact FVM and derived some high accuracy post-processing formulas in [17]. Li proved the optimal L^2 error estimate for bilinear and biquadratic FVMs with smooth coefficients under a mesh restriction of h^2 -parallelogram, respectively ([18, 19]). In addition, there are many works for the error estimates of finite element method to solve diffusion problems with smooth coefficients, for example, Nie derived the optimal L^2 error estimate of linear finite element scheme for the diffusion problem with nonlocal boundary in [20]. However, there have been few strict theoretical analyses of the global error estimation for Eulerian FVMs to solve diffusion problems with strong discontinuity and multiple material cells.

In the late 1980's, M.H. Frese presented a finite volume method for diffusion equations in 2D magnetohydrodynamic problems on the quadrilateral grids ([21]). The corresponding software packages named MACH2 and MACH3 for 2D and 3D problems have been successfully developed, respectively, by Philips Lab/WSP, Kirtland Air Force Base ([21, 22]), and widely used in the numerical simulations of liner implosion system, plasma thrusters and so on (see, e.g. [23, 24, 25, 26, 27]). For simplicity of presentation, we denote this finite volume method as MACH FVM. However, to our knowledge, we have not found the strict error theories of it for diffusion problems with strong discontinuity and multiple material cells on the Eulerian grids, which urges us to study it.

In this paper, we present a vertex-centered MACH-like FVM on the quadrilateral grids for stationary diffusion problems with interface. In particular, for the square grids, this kind of nine-point scheme is degenerated into a five-point scheme. It is worth pointing out that many classical nine-point FVMs (see, for example, [28]) are always degenerated into a five-point stencil which looks like “+”, while the stencil of the five-point MACH-like scheme looks like “×”. Therefore, this adds an extra difficulty for error estimates. The other important work of this paper is that we present the strict theoretical analysis for the five-point MACH-like scheme. It is divided into two parts. On the one hand, we discuss the local truncation error. The main difficulty results from the discontinuous coefficients of the interface. By the homogeneous jump conditions on the interface and Taylor expansions, we derive that the local truncation error of the interior nodes adjacent to the interface is $O(h)$. Furthermore, we get the local truncation error $O(h^2)$ under the Assumption I (i.e., we use harmonic average method for the diffusion coefficients and the second derivative function with respect to the tangent direction of the exact solution is equal to zero on the interface). On the other hand, we focus on the global error estimates. Firstly, the error difference equations are decomposed into two relatively simple ones. Then, by using the discrete sine transform ([29, 30]) and combining with some analytical techniques, we convert these 2D difference equations into two kinds of one-dimensional (1D) difference equations, and the estimates of these difference equations are deduced. Hence, we demonstrate that the global error estimation of the five-point scheme is $O(h|\ln h|)$ in the maximum norm. Furthermore, under the Assumption I, we obtain the asymptotic optimal error estimate $O(h^2|\ln h|)$. In addition, we investigate the approximation of the five-point MACH-like scheme by several typical numerical examples. Numerical results are carried on to confirm the theoretical ones.

The paper is organized as follows. In Section 2, we describe the construction of the MACH-like

finite volume method on an arbitrary quadrilateral grid, and present a five-point MACH-like scheme for special. Section 3 presents the local truncation error for the five-point scheme. Then the global error estimation in the maximum norm is derived in Section 4. After that, the accuracy of the five-point MACH-like scheme is verified numerically. In Section 6 we draw some conclusions on our works.

2. Model problem and finite volume scheme

We consider the following interface problem

$$\begin{cases} -\nabla \cdot (\kappa \nabla u) + u = f, & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega, \end{cases} \quad (2.1)$$

where $\Omega \in \mathbb{R}^2$ is a bounded polygonal domain with boundary $\partial\Omega$, f is a given function and the diffusion coefficient κ is positive and piecewise constant on polygonal subdomains of Ω with possible large discontinuities across subdomain boundaries which is simply interface for short. Let $\kappa = \kappa_i > 0$ in D_i , for $i = 1, 2, \dots, J$. Here $\{D_i\}_{i=1}^J$ is a partition of Ω , where D_i is an open polygonal subdomain.

Denote $\Gamma_{i,j} = \bar{D}_i \cap \bar{D}_j (i \neq j)$ and $\Gamma = \cup_{i,j=1}^J \Gamma_{i,j}$. We assume that

$$[u] = [\kappa \frac{\partial u}{\partial \vec{n}}] = 0, \quad \text{on } \Gamma, \quad (2.2)$$

where \vec{n} is the unit outward normal vector on Γ and $[\zeta](\zeta = u, \kappa \frac{\partial u}{\partial \vec{n}})$ denotes the difference of the right and left limits of ζ at any point of Γ .

Let \mathcal{Q}_h be a structured quadrilateral partition of Ω , and $X = \{X_{i,j} = (x_i, y_j), i = 0, 1, \dots, N_x, j = 0, 1, \dots, N_y\}$, where N_x, N_y are given positive integers.

Next, We derive a kind of FVM with vertex unknowns for solving (2.1) and (2.2), which is based on the method of Michael H. Frese in [21]. We denote it as MACH-like FVM for convenience.

For any given interior node $X_{i,j}, i = 1, 2, \dots, N_x - 1, j = 1, 2, \dots, N_y - 1$ (see Fig. 1). We denote open area $Q_l (l = 1, 2, 3, 4)$ as the l -th quadrilateral element adjacent to $X_{i,j}$, where $O_l, l = 1, 2, 3, 4$ are the midpoints of the grid sides $X_{i,j-1}X_{i,j}, X_{i+1,j}X_{i,j}, X_{i,j+1}X_{i,j}$ and $X_{i-1,j}X_{i,j}$, respectively. The quadrangle $X_{i,j}O_{l-1}A_lO_l$ is a parallelogram, where $A_l \in Q_l$. Thus, the control volume of $X_{i,j}$ is defined as the octagon area which is surrounded by $A_l, O_l (l = 1, 2, 3, 4)$. We denote it briefly by $V_{i,j}$.

Fig. 1 (a) and (b) show the typical cases which have an interface in $Q_l (l = 1, 2, 3, 4)$ or not.

Integrating (2.1) over $V_{i,j}$, the equation (2.1) leads to

$$\int_{V_{i,j}} (-\nabla \cdot (\kappa \nabla u) + u) dX = \int_{V_{i,j}} f dX.$$

By the divergence theorem and the continuity condition of the flux $\kappa \frac{\partial u}{\partial \vec{n}}$, we get

$$-\int_{\partial V_{i,j}} \kappa \frac{\partial u}{\partial \vec{n}} dS + \int_{V_{i,j}} u dX = \int_{V_{i,j}} f dX, \quad (2.3)$$

where $\partial V_{i,j}$ is the boundary of $V_{i,j}$ and \vec{n} is the unit outward normal vector of $\partial V_{i,j}$.

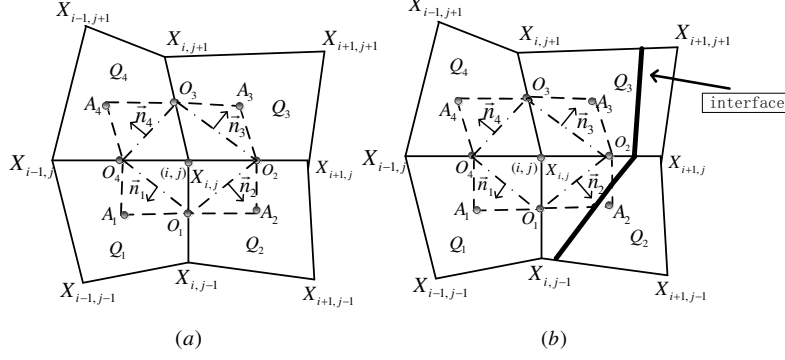


Figure 1: (a) No interface in Q_l ($l = 1, 2, 3, 4$). (b) Only one interface in Q_2 and Q_3 .

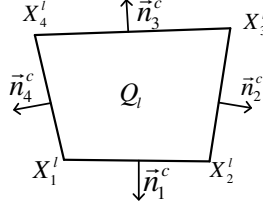


Figure 2: Illustration for Q_l .

Denote \vec{n}_l as the unit outward normal vector on the side $O_{l-1}O_l$ for $l = 1, 2, 3, 4$ (see Fig. 1). Combining (2.3) and the rectangular integral formula, we can obtain

$$\begin{aligned} & -\kappa_1^*(\nabla u)|_{Q_1} \cdot \vec{n}_1 |O_4O_1| - \kappa_2^*(\nabla u)|_{Q_2} \cdot \vec{n}_2 |O_1O_2| \\ & -\kappa_3^*(\nabla u)|_{Q_3} \cdot \vec{n}_3 |O_2O_3| - \kappa_4^*(\nabla u)|_{Q_4} \cdot \vec{n}_4 |O_3O_4| + |V_{i,j}|u_{i,j} \approx \int_{V_{i,j}} f dX, \end{aligned} \quad (2.4)$$

where κ_l^* is the averaging of κ over Q_l and $(\nabla u)|_{Q_l}$ is the integral average value over Q_l .

Hence, from (2.4), we know that it is an urgent task to get the approximate calculation formula of $(\nabla u)|_{Q_l}$ for $l = 1, 2, 3, 4$.

Any open area Q_l is shown in Fig. 2, where X_m^l ($m = 1, 2, 3, 4$) is the vertex of Q_l , \vec{n}_m^c ($m = 1, 2, 3, 4$) is the unit outward normal vector on the side $X_m^l X_{m+1}^l$ ($m = 1, 2, 3, 4$, $X_5^l = X_1^l$).

Motivated by the finite volume method in M. H. Frese [21], we can obtain

$$\begin{aligned} (\nabla u)|_{Q_l} & \approx \frac{1}{|Q_l|} \left[\frac{1}{2} (u(X_1^l) + u(X_2^l)) |X_1^l X_2^l| \vec{n}_1^c + \frac{1}{2} (u(X_2^l) + u(X_3^l)) |X_2^l X_3^l| \vec{n}_2^c \right. \\ & \left. + \frac{1}{2} (u(X_3^l) + u(X_4^l)) |X_3^l X_4^l| \vec{n}_3^c + \frac{1}{2} (u(X_4^l) + u(X_1^l)) |X_4^l X_1^l| \vec{n}_4^c \right], \end{aligned} \quad (2.5)$$

where $|\zeta|$ is the length or area of ζ .

The option of κ_l^* ($l = 1, 2, 3, 4$) is another critical element to construct the MACH-like FVM. For any given Q_l ($l = 1, 2, 3, 4$), if Q_l is a single material area (see Fig. 1(a) Q_k ($k = 1, 2, 3, 4$) as an example), without loss of generality, assuming that $Q_l \subset D_m$ ($1 \leq m \leq J$), then it follows that $\kappa_l^* = \kappa_m$. If Q_l is a multiple material area (see, for example, Q_2, Q_3 in Fig. 1(b)), then κ_l^* must

be obtained from an appropriate average method, such as arithmetic average method, harmonic average method and so on.

Substituting (2.5) into (2.4), and combining with the proper κ_l^* ($l = 1, 2, 3, 4$), we get the MACH-like FVM of (2.1) and (2.2) at $X_{i,j}$ ($i = 1, 2, \dots, N_x - 1, j = 1, 2, \dots, N_y - 1$).

Remark 2.1. *The MACH-like FVM is always a compact nine-point stencil.*

In particular, we consider a uniform quadrilateral grid in $\Omega = (a, b) \times (c, d)$, where

$$\begin{aligned} h_x &= (b - a)/N_x, \quad h_y = (d - c)/N_y, \\ (x_i, y_j) &= (a + ih_x, c + jh_y), \quad i = 0, 1, \dots, N_x, j = 0, 1, \dots, N_y. \end{aligned}$$

The MACH-like FVM at any interior node $X_{i,j}$ express as follows

$$\begin{aligned} &a_1^{i,j} u_{i-1,j-1} + a_2^{i,j} u_{i,j-1} + a_3^{i,j} u_{i+1,j-1} + a_4^{i,j} u_{i-1,j} + a_5^{i,j} u_{i,j} \\ &+ a_6^{i,j} u_{i+1,j} + a_7^{i,j} u_{i-1,j+1} + a_8^{i,j} u_{i,j+1} + a_9^{i,j} u_{i+1,j+1} = h_x h_y f_{i,j}, \end{aligned} \quad (2.6)$$

where

$$\left\{ \begin{aligned} a_1^{i,j} &= -\frac{1}{4} \left(\frac{h_y}{h_x} + \frac{h_x}{h_y} \right) \kappa_1^*, \\ a_2^{i,j} &= -\frac{1}{4} \left(-\frac{h_y}{h_x} + \frac{h_x}{h_y} \right) (\kappa_1^* + \kappa_2^*), \\ a_3^{i,j} &= -\frac{1}{4} \left(\frac{h_y}{h_x} + \frac{h_x}{h_y} \right) \kappa_2^*, \\ a_4^{i,j} &= -\frac{1}{4} \left(\frac{h_y}{h_x} - \frac{h_x}{h_y} \right) (\kappa_1^* + \kappa_4^*), \\ a_5^{i,j} &= \frac{1}{4} \left(\frac{h_y}{h_x} + \frac{h_x}{h_y} \right) (\kappa_1^* + \kappa_2^* + \kappa_3^* + \kappa_4^*) + h_x h_y, \\ a_6^{i,j} &= -\frac{1}{4} \left(\frac{h_y}{h_x} - \frac{h_x}{h_y} \right) (\kappa_2^* + \kappa_3^*), \\ a_7^{i,j} &= -\frac{1}{4} \left(\frac{h_y}{h_x} + \frac{h_x}{h_y} \right) \kappa_4^*, \\ a_8^{i,j} &= -\frac{1}{4} \left(-\frac{h_y}{h_x} + \frac{h_x}{h_y} \right) (\kappa_3^* + \kappa_4^*), \\ a_9^{i,j} &= -\frac{1}{4} \left(\frac{h_y}{h_x} + \frac{h_x}{h_y} \right) \kappa_3^*, \\ f_{i,j} &= \frac{1}{h_x h_y} \int_{V_{i,j}} f dX. \end{aligned} \right. \quad (2.7)$$

Currently, MACH scheme is successfully applied in the numerical simulation of liner implosion system, plasma thrusters and so on (see, e.g. [23, 25, 26]). However, there is few error theories about this scheme, and it is always for the local truncation error. Especially, the strict error theories for the stationary diffusion problems with multiple material cells on Eulerian grids haven't been seen yet. In the following sections, we will establish a rigorous theoretical analysis of the MACH-like scheme, which is constructed for a simplified model of (2.1) and (2.2).

This simplified model only considers two subdomains. Suppose that $\Omega = (0, 1) \times (0, 1) = D_1 \cup D_2 \cup \Gamma$ and the diffusion coefficient

$$\kappa = \begin{cases} \kappa^- \geq 1, & \text{in } D_1, \\ 1, & \text{in } D_2, \end{cases} \quad (2.8)$$

where $D_1 = (0, \frac{1}{2}) \times (0, 1)$, $D_2 = (\frac{1}{2}, 1) \times (0, 1)$ and κ^- is a positive constant.

In addition, we assume that \mathcal{Q}_h is a square mesh of Ω , where $N_x = N_y = N = 2M + 1$ ($M \in \mathbb{Z}^+$), $h_x = h_y = h$, introduce the grid nodes and the interior nodes indicated set $S = \{(i, j), i, j = 0, 1, \dots, N\}$ and $S_0 = \{(i, j), i, j = 1, 2, \dots, N - 1\}$. Let the indicated set of the interior nodes which are adjacent to the interface or not be

$$S_B = S_1^B \cup S_2^B \quad \text{and} \quad S_I = S_1^I \cup S_2^I, \quad (2.9)$$

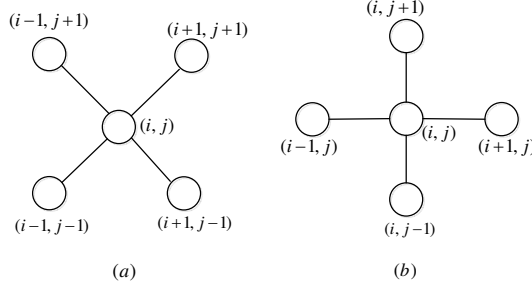


Figure 3: (a) “x” type five-point stencil. (b) “+” type five-point stencil.

respectively, where

$$\begin{aligned}
S_1^I &= \{(i, j), i = 1, 2, \dots, M-1, j = 1, 2, \dots, N-1\}, \\
S_2^I &= \{(i, j), i = M+2, M+3, \dots, N-1, j = 1, 2, \dots, N-1\}, \\
S_1^B &= \{(i, j), i = M, j = 1, 2, \dots, N-1\}, \\
S_2^B &= \{(i, j), i = M+1, j = 1, 2, \dots, N-1\}.
\end{aligned}$$

Denote κ^* as the average value of the diffusion coefficient in the multiple material cells. From (2.6) and (2.7), we can obtain the five-point MACH-like scheme of the simplified model as follows

$$\begin{cases} L_h u_{i,j} = h^2 f_{i,j}, & (i, j) \in S_0, \\ u_{0,j} = u_{N,j} = u_{i,0} = u_{i,N} = 0, & (i, j) \in S, \end{cases} \quad (2.10)$$

where

$$L_h u_{i,j} = a_1^{i,j} u_{i-1,j-1} + a_2^{i,j} u_{i+1,j-1} + a_3^{i,j} u_{i,j} + a_4^{i,j} u_{i-1,j+1} + a_5^{i,j} u_{i+1,j+1}, \quad f_{i,j} = f(x_i, y_j), \quad (2.11)$$

and

$$\begin{cases} a_1^{i,j} = a_2^{i,j} = a_4^{i,j} = a_5^{i,j} = -\frac{1}{2}\kappa^-, & a_3^{i,j} = 2\kappa^- + h^2, & (i, j) \in S_1^I, \\ a_1^{i,j} = a_4^{i,j} = -\frac{1}{2}\kappa^-, & a_2^{i,j} = a_5^{i,j} = -\frac{1}{2}\kappa^*, & a_3^{i,j} = \kappa^- + \kappa^* + h^2, & (i, j) \in S_1^B, \\ a_1^{i,j} = a_4^{i,j} = -\frac{1}{2}\kappa^*, & a_2^{i,j} = a_5^{i,j} = -\frac{1}{2}, & a_3^{i,j} = \kappa^* + 1 + h^2, & (i, j) \in S_2^B, \\ a_1^{i,j} = a_2^{i,j} = a_4^{i,j} = a_5^{i,j} = -\frac{1}{2}, & a_3^{i,j} = 2 + h^2, & & (i, j) \in S_2^I. \end{cases}$$

The stencil of this five-point scheme which looks like “x” is shown in Fig. 3 (a). It is worth pointing out that many frequently-used nine-point FVMs, such as [28], are degenerated into a five-point stencil which looks like “+” (see Fig. 3(b)). Hence, this difference will bring some new difficulties to the following error analysis.

In the rest of this paper, we will present the rigorous error analysis and corresponding numerical experiments for the five-point MACH-like scheme (2.10).

3. Local truncation error estimation

Let

$$R_{i,j} = L_h u(x_i, y_j) - L_h u_{i,j} = L_h u(x_i, y_j) - h^2 f_{i,j}, \quad (i, j) \in S_0 \quad (3.1)$$

be the local truncation error of (2.10) at $X_{i,j}$, and denote $x_{i+\frac{1}{2}} = x_i + \frac{h}{2}$ ($i = 0, 1, \dots, N-1$), $\delta_{l,m}$ as the Kronecker delta.

Define a function space via

$$W = \{w \mid w|_{\bar{D}_i} \in C^6(\bar{D}_i), i = 1, 2\}.$$

In this section, we will present the local truncation error of (2.10).

Theorem 3.1. *If $u \in W$, then the following holds.*

1) *For any interior node which is non adjacent to the interface, we have*

$$R_{i,j} = (\kappa^-)^{\delta_{1,k}} C^{(4)}(x_i, y_j) h^4 + O(h^6), \quad (i, j) \in S_k^I, \quad k = 1, 2, \quad (3.2)$$

where

$$C^{(4)}(x, y) = -\frac{1}{12} \left(\frac{\partial^4 u}{\partial y^4} + \frac{\partial^4 u}{\partial x^4} + 6 \frac{\partial^4 u}{\partial y^2 \partial x^2} \right) (x, y). \quad (3.3)$$

2) *For any interior node which is adjacent to the interface, we have*

$$R_{i,j} = C_i^{(1)}(y_j) h + C_i^{(2)}(y_j) h^2 + C_i^{(3)}(y_j) h^3 + O(h^4), \quad (i, j) \in S_B, \quad (3.4)$$

where

$$C_M^{(1)}(y) = (\kappa^- - \frac{\kappa^*}{2} - \frac{\kappa^* \kappa^-}{2}) \left(\frac{\partial u}{\partial x} \right)^-(y), \quad C_{M+1}^{(1)}(y) = -C_M^{(1)}(y), \quad (3.5)$$

$$C_M^{(2)}(y) = \frac{\kappa^*}{8} \left[\frac{\partial^2 u}{\partial x^2} \right](y) + \frac{1}{2} \left(\kappa^- \left(\frac{\partial^2 u}{\partial y^2} \right)^- - \kappa^* \left(\frac{\partial^2 u}{\partial y^2} \right)^+ \right) (y), \quad (3.6)$$

$$C_{M+1}^{(2)}(y) = -\frac{\kappa^*}{8} \left[\frac{\partial^2 u}{\partial x^2} \right](y) + \frac{1}{2} \left(\left(\frac{\partial^2 u}{\partial y^2} \right)^+ - \kappa^* \left(\frac{\partial^2 u}{\partial y^2} \right)^- \right) (y), \quad (3.7)$$

$$C_M^{(3)}(y) = \left(\frac{\kappa^-}{24} \left(\frac{\partial^3 u}{\partial x^3} \right)^- - \frac{\kappa^*}{24} \left\{ \frac{\partial^3 u}{\partial x^3} \right\} + \frac{1}{4} \left(\kappa^- \left(\frac{\partial^3 u}{\partial y^2 \partial x} \right)^- - \kappa^* \left(\frac{\partial^3 u}{\partial y^2 \partial x} \right)^+ \right) \right) (y), \quad (3.8)$$

$$C_{M+1}^{(3)}(y) = \left(-\frac{1}{24} \left(\frac{\partial^3 u}{\partial x^3} \right)^+ + \frac{\kappa^*}{24} \left\{ \frac{\partial^3 u}{\partial x^3} \right\} + \frac{1}{4} \left(-\left(\frac{\partial^3 u}{\partial y^2 \partial x} \right)^+ + \kappa^* \left(\frac{\partial^3 u}{\partial y^2 \partial x} \right)^- \right) \right) (y), \quad (3.9)$$

$$\text{and } (\zeta)^\pm(y) := \lim_{x \rightarrow x_{M+\frac{1}{2}}^\pm} \zeta(x, y), \quad [\zeta] := (\zeta)^- - (\zeta)^+, \quad \{\zeta\} := \frac{1}{2}((\zeta)^- + (\zeta)^+).$$

Proof. We only prove for $i = M$ and $j = 1, 2, \dots, N-1$, the remainder of the argument is analogous to it.

By (2.11) and (3.1), we have

$$\begin{aligned} R_{M,j} &= -\frac{1}{2} \kappa^- [u(x_{M-1}, y_{j-1}) + u(x_{M-1}, y_{j+1})] - \frac{1}{2} \kappa^* [u(x_{M+1}, y_{j-1}) + u(x_{M+1}, y_{j+1})] \\ &\quad + (\kappa^- + \kappa^* + h^2) u(x_M, y_j) - h^2 f(x_M, y_j). \end{aligned} \quad (3.10)$$

Substituting the Taylor series expansions of $u(x_{M\pm 1}, y_{j-1})$ and $u(x_{M\pm 1}, y_{j+1})$ at $(x_{M\pm 1}, y_j)$ into (3.10), we have

$$\begin{aligned} R_{M,j} &= \kappa^- [u(x_M, y_j) - u(x_{M-1}, y_j)] + \kappa^* [u(x_M, y_j) - u(x_{M+1}, y_j)] \\ &\quad - \frac{1}{2} h^2 \left[\kappa^- \frac{\partial^2 u}{\partial y^2}(x_{M-1}, y_j) + \kappa^* \frac{\partial^2 u}{\partial y^2}(x_{M+1}, y_j) \right] \\ &\quad + h^2 u(x_M, y_j) - h^2 f(x_M, y_j) + O(h^4). \end{aligned} \quad (3.11)$$

Further, substituting the Taylor series expansions of $u(x_M, y_j)$, $u(x_{M\pm 1}, y_j)$, $f(x_M, y_j)$ and $\frac{\partial^2 u}{\partial y^2}(x_{M\pm 1}, y_j)$ at $(x_{M+\frac{1}{2}}, y_j)$ into (3.11), and using $[u]|_\Gamma = 0$, equation (2.1) and $[\kappa \frac{\partial u}{\partial n}]|_\Gamma = 0$, $R_{M,j}$ is rewritten as

$$\begin{aligned}
R_{M,j} &= h(\kappa^- - \frac{\kappa^*}{2} - \frac{\kappa^* \kappa^-}{2})(\frac{\partial u}{\partial x})^-(y_j) \\
&\quad + h^2(\frac{\kappa^*}{8}[\frac{\partial^2 u}{\partial x^2}] + \frac{1}{2}(\kappa^-(\frac{\partial^2 u}{\partial y^2})^- - \kappa^*(\frac{\partial^2 u}{\partial y^2})^+))(y_j) \\
&\quad + h^3(\frac{\kappa^-}{24}(\frac{\partial^3 u}{\partial x^3})^- - \frac{\kappa^*}{24}\{\frac{\partial^3 u}{\partial x^3}\} + \frac{1}{4}(\kappa^-(\frac{\partial^3 u}{\partial y^2 \partial x})^- - \kappa^*(\frac{\partial^3 u}{\partial y^2 \partial x})^+))(y_j) + O(h^4) \\
&=: C_M^{(1)}(y_j)h + C_M^{(2)}(y_j)h^2 + C_M^{(3)}(y_j)h^3 + O(h^4),
\end{aligned} \tag{3.12}$$

where $C_M^{(1)}(y)$, $C_M^{(2)}(y)$ and $C_M^{(3)}(y)$ are defined by (3.5), (3.6) and (3.8), respectively.

This finishes the proof. \square

For a lot of diffusion problems with interface (such as [31]), the changing of the solution along the normal direction (here for the x direction) of the interfaces is far more quickly than the tangential direction (here for the y direction). As a result, we suppose that

$$\lim_{x \rightarrow x_{M+\frac{1}{2}}^-} \frac{\partial^2 u}{\partial y^2} = \lim_{x \rightarrow x_{M+\frac{1}{2}}^+} \frac{\partial^2 u}{\partial y^2} = 0. \tag{3.13}$$

Under the above assumption, we can obtain a corollary from (3.4) as follows.

Corollary 3.1. *Let $u \in W$ and suppose that $\kappa^* = \frac{2\kappa^-}{\kappa^-+1}$ and (3.13) are hold, then*

$$R_{i,j} = \tilde{C}_i^{(2)}(y_j)h^2 + C_i^{(3)}(y_j)h^3 + O(h^4), \quad (i, j) \in S_B, \tag{3.14}$$

where

$$\tilde{C}_M^{(2)}(y) = \frac{\kappa^*}{8}[\frac{\partial^2 u}{\partial x^2}](y) \quad \text{and} \quad \tilde{C}_{M+1}^{(2)}(y) = -\tilde{C}_M^{(2)}(y). \tag{3.15}$$

For simplicity, we denote the assumptions $\kappa^* = \frac{2\kappa^-}{\kappa^-+1}$ and (3.13) as Assumption **I**.

4. Global error estimation

In this section, we will investigate the global error estimates for the five-point MACH-like scheme. Defined

$$\vec{e} = (e_{0,0}, \dots, e_{0,N}, \dots, e_{N,0}, \dots, e_{N,N}) \tag{4.1}$$

as the error vector, where

$$e_{i,j} = u(x_i, y_j) - u_{i,j}. \tag{4.2}$$

Using (4.1), (4.2) and (2.10), we can obtain that \vec{e} satisfies the following 2D difference equations

$$\begin{cases} L_h e_{i,j} = R_{i,j}, & (i, j) \in S_0, \\ e_{i,0} = e_{i,N} = e_{0,j} = e_{N,j} = 0, & (i, j) \in S, \end{cases} \tag{4.3}$$

where $R_{i,j}$ for $(i,j) \in S_0$ are given by (3.2) and (3.4).

Let S_B be the indicated set of the interior nodes adjacent to the interface(given by (2.9)), denote $e_{i,j} = 0 (\forall (i,j) \in S_B)$ in (4.3), and we have

$$\begin{cases} -\frac{1}{2}\kappa^-(e_{i-1,j-1}^{(1)} + e_{i-1,j+1}^{(1)}) + (2\kappa^- + h^2)e_{i,j}^{(1)} - \frac{1}{2}\kappa^-(e_{i+1,j-1}^{(1)} + e_{i+1,j+1}^{(1)}) = R_{i,j}, & (i,j) \in S_1^I, \\ -\frac{1}{2}(e_{i-1,j-1}^{(1)} + e_{i-1,j+1}^{(1)}) + (2 + h^2)e_{i,j}^{(1)} - \frac{1}{2}(e_{i+1,j-1}^{(1)} + e_{i+1,j+1}^{(1)}) = R_{i,j}, & (i,j) \in S_2^I, \\ e_{i,j}^{(1)} = 0, & (i,j) \in S_B, \\ e_{0,j}^{(1)} = e_{N,j}^{(1)} = e_{i,0}^{(1)} = e_{i,N}^{(1)} = 0, & (i,j) \in S. \end{cases} \quad (4.4)$$

Let $\vec{e}^{(1)} = (e_{0,0}^{(1)}, \dots, e_{0,N}^{(1)}, \dots, e_{N,0}^{(1)}, \dots, e_{N,N}^{(1)})$ be the solution vector of (4.4) and the error vector \vec{e} can be decomposed as

$$\vec{e} = \vec{e}^{(1)} + \vec{e}^{(2)}. \quad (4.5)$$

By (4.3), (4.4) and (4.5), it is easy to check that $\vec{e}^{(2)} := (e_{0,0}^{(2)}, \dots, e_{0,N}^{(2)}, \dots, e_{N,0}^{(2)}, \dots, e_{N,N}^{(2)})$ satisfies the following difference equations

$$\begin{cases} -\frac{1}{2}\kappa^-(e_{i-1,j-1}^{(2)} + e_{i-1,j+1}^{(2)}) + (2\kappa^- + h^2)e_{i,j}^{(2)} - \frac{1}{2}\kappa^-(e_{i+1,j-1}^{(2)} + e_{i+1,j+1}^{(2)}) = 0, & (i,j) \in S_1^I, \\ -\frac{1}{2}\kappa^-(e_{i-1,j-1}^{(2)} + e_{i-1,j+1}^{(2)}) + (\kappa^- + \kappa^* + h^2)e_{i,j}^{(2)} - \frac{1}{2}\kappa^*(e_{i+1,j-1}^{(2)} + e_{i+1,j+1}^{(2)}) = \gamma_{M,j}, & (i,j) \in S_1^B, \\ -\frac{1}{2}\kappa^*(e_{i-1,j-1}^{(2)} + e_{i-1,j+1}^{(2)}) + (\kappa^* + 1 + h^2)e_{i,j}^{(2)} - \frac{1}{2}(e_{i+1,j-1}^{(2)} + e_{i+1,j+1}^{(2)}) = \gamma_{M+1,j}, & (i,j) \in S_2^B, \\ -\frac{1}{2}(e_{i-1,j-1}^{(2)} + e_{i-1,j+1}^{(2)}) + (2 + h^2)e_{i,j}^{(2)} - \frac{1}{2}(e_{i+1,j-1}^{(2)} + e_{i+1,j+1}^{(2)}) = 0, & (i,j) \in S_2^I, \\ e_{0,j}^{(2)} = e_{N,j}^{(2)} = e_{i,0}^{(2)} = e_{i,N}^{(2)} = 0, & (i,j) \in S, \end{cases} \quad (4.6)$$

where

$$\gamma_{M,j} = R_{M,j} + \frac{1}{2}\kappa^-(e_{M-1,j-1}^{(1)} + e_{M-1,j+1}^{(1)}), \quad \gamma_{M+1,j} = R_{M+1,j} + \frac{1}{2}(e_{M+2,j-1}^{(1)} + e_{M+2,j+1}^{(1)}).$$

Then, we will employ the dimension reduction techniques (convert the 2D problems into the 1D problems) to estimate $\vec{e}^{(l)} (l = 1, 2)$. Therefore, for any given $i = 1, 2, \dots, N-1$, $l = 1, 2$, introduce the discrete sine transform (see, e.g. [29]) for sequence $\{e_{i,j}^{(l)}\}_{j=1}^{N-1}$ and $\{R_{i,j}\}_{j=1}^{N-1}$ as

$$\bar{e}_{i,k}^{(l)} = \sqrt{2h} \sum_{j=1}^{N-1} e_{i,j}^{(l)} \sin(j\pi y_k), \quad \bar{R}_{i,k} = \sqrt{2h} \sum_{j=1}^{N-1} R_{i,j} \sin(j\pi y_k), \quad k = 1, 2, \dots, N-1. \quad (4.7)$$

Using

$$\sin(k\pi y_j) = \sin(j\pi y_k), \quad (k,j) \in S_0,$$

and

$$\sum_{k=1}^{N-1} \sin(k\pi y_l) \sin(k\pi y_m) = \frac{1}{2h} \delta_{l,m},$$

one can easily confirm

$$e_{i,j}^{(l)} = \sqrt{2h} \sum_{k=1}^{N-1} \bar{e}_{i,k}^{(l)} \sin(k\pi y_j), \quad R_{i,j} = \sqrt{2h} \sum_{k=1}^{N-1} \bar{R}_{i,k} \sin(k\pi y_j), \quad j = 1, 2, \dots, N-1, \quad (4.8)$$

and where the sequences $\{e_{i,j}^{(l)}\}_{j=1}^{N-1}$ and $\{R_{i,j}\}_{j=1}^{N-1}$ are called as the inverse discrete sine transform of $\{\bar{e}_{i,k}^{(l)}\}_{k=1}^{N-1}$ and $\{\bar{R}_{i,k}\}_{k=1}^{N-1}$, respectively.

Let

$$\bar{e}_k^{(l)} = (\bar{e}_{0,k}^{(l)}, \dots, \bar{e}_{N,k}^{(l)}), \quad k = 1, 2, \dots, N-1, l = 1, 2, \quad (4.9)$$

where $\bar{e}_{i,k}^{(l)} (i = 1, 2, \dots, N-1)$ is defined by (4.7), and

$$\bar{e}_{0,k}^{(l)} = \bar{e}_{N,k}^{(l)} = 0. \quad (4.10)$$

Using (4.7) and (4.8), we can convert the estimate problems of $\bar{e}^{(l)} (l = 1, 2)$ into $\bar{e}_k^{(l)} (k = 1, 2, \dots, N-1)$.

Denote the maximum norm of any vector $\vec{\zeta} = (\zeta_1, \dots, \zeta_n)$ as $\|\vec{\zeta}\|_\infty$, and it is calculated as $\max_i |\zeta_i|$. We will present the estimation of $\|\bar{e}_k^{(1)}\|_\infty (k = 1, 2, \dots, N-1)$ firstly. It is easy to verify that $\bar{e}_k^{(1)} (k = 1, 2, \dots, N-1)$ satisfies the following 1D difference equations

$$\begin{cases} -\kappa^- \bar{e}_{i-1,k}^{(1)} \cos(k\pi h) + (2\kappa^- + h^2) \bar{e}_{i,k}^{(1)} - \kappa^- \bar{e}_{i+1,k}^{(1)} \cos(k\pi h) = \bar{R}_{i,k}, & i = 1, 2, \dots, M-1, \\ -\bar{e}_{i-1,k}^{(1)} \cos(k\pi h) + (2 + h^2) \bar{e}_{i,k}^{(1)} - \bar{e}_{i+1,k}^{(1)} \cos(k\pi h) = \bar{R}_{i,k}, & i = M+2, M+3, \dots, N-1, \\ \bar{e}_{M,k}^{(1)} = \bar{e}_{M+1,k}^{(1)} = 0, \quad \bar{e}_{0,k}^{(1)} = \bar{e}_{N,k}^{(1)} = 0. \end{cases} \quad (4.11)$$

As a matter of fact, using (4.7) and noting that $e_{i,j}^{(1)} = 0 (\forall (i, j) \in S_B)$, we can get $\bar{e}_{M,k}^{(1)} = \bar{e}_{M+1,k}^{(1)} = 0$. Hence, from (4.10), it follows that we only need to demonstrate the former $N-3$ equations of (4.11) are established.

Substituting (4.8) into the part of (4.4) where $(i, j) \in S_1^I$, we get

$$\begin{aligned} & -\frac{1}{2} \kappa^- \sqrt{2h} \sum_{k=1}^{N-1} \bar{e}_{i-1,k}^{(1)} [\sin(k\pi y_{j-1}) + \sin(k\pi y_{j+1})] + (2\kappa^- + h^2) \sqrt{2h} \sum_{k=1}^{N-1} \bar{e}_{i,k}^{(1)} \sin(k\pi y_j) \\ & -\frac{1}{2} \kappa^- \sqrt{2h} \sum_{k=1}^{N-1} \bar{e}_{i+1,k}^{(1)} [\sin(k\pi y_{j-1}) + \sin(k\pi y_{j+1})] = \sqrt{2h} \sum_{k=1}^{N-1} \bar{R}_{i,k} \sin(k\pi y_j), \quad j = 1, 2, \dots, N-1. \end{aligned} \quad (4.12)$$

Owing to the characterization property of the discrete sine function $\sin(k\pi y_j) (k, j = 1, 2, \dots, N)$ that

$$\sin(k\pi y_{j-1}) + \sin(k\pi y_{j+1}) = 2 \sin(k\pi y_j) \cos(k\pi h),$$

equation (4.12) can be further represented as

$$\begin{aligned} & \sqrt{2h} \sum_{k=1}^{N-1} [-\kappa^- \bar{e}_{i-1,k}^{(1)} \cos(k\pi h) + (2\kappa^- + h^2) \bar{e}_{i,k}^{(1)} - \kappa^- \bar{e}_{i+1,k}^{(1)} \cos(k\pi h) - \bar{R}_{i,k}] \sin(k\pi y_j) = 0, \\ & j = 1, 2, \dots, N-1. \end{aligned}$$

Therefore, we have

$$-\kappa^- \bar{e}_{i-1,k}^{(1)} \cos(k\pi h) + (2\kappa^- + h^2) \bar{e}_{i,k}^{(1)} - \kappa^- \bar{e}_{i+1,k}^{(1)} \cos(k\pi h) - \bar{R}_{i,k} = 0, \quad k = 1, 2, \dots, N-1.$$

Namely, the former $M-1$ equations of (4.11) are established.

The rest of the proof is similar to before. The proof is completed.

In order to estimate $\|\bar{e}_k^{(1)}\|_\infty (k = 1, 2, \dots, N-1)$ by using the difference equations (4.11), we need to introduce several lemmas as follows.

Lemma 4.1. Let $\varphi_j := \varphi(y_j)(j = 1, 2, \dots, N-1)$, and assume that $\varphi(y) \in C^2([0, 1])$, then

$$\left| \sum_{j=1}^{N-1} \varphi_j \sin(j\pi y_k) \right| \lesssim \frac{\|\varphi\|_{2,\infty}}{\sin \frac{k\pi h}{2}}. \quad (4.13)$$

Proof. Since

$$\begin{aligned} \sum_{j=2}^{N-2} (\varphi_{j-1} + \varphi_{j+1}) \sin(j\pi y_k) &= \sum_{j=1}^{N-3} \varphi_j \sin((j+1)\pi y_k) + \sum_{j=3}^{N-1} \varphi_j \sin((j-1)\pi y_k) \\ &= 2 \cos(\pi y_k) \sum_{j=1}^{N-1} \varphi_j \sin(j\pi y_k) - \sin(\pi y_k)((-1)^{k-1} \varphi_{N-2} - \varphi_2), \end{aligned}$$

we have

$$\begin{aligned} (\cos(\pi y_k) - 1) \sum_{j=1}^{N-1} \varphi_j \sin(j\pi y_k) &= \sum_{j=2}^{N-2} \left(\frac{\varphi_{j-1} + \varphi_{j+1}}{2} - \varphi_j \right) \sin(j\pi y_k) \\ &\quad + \sin(\pi y_k)((-1)^{k-1} \left(\frac{1}{2} \varphi_{N-2} - \varphi_{N-1} \right) + \left(\frac{1}{2} \varphi_2 - \varphi_1 \right)). \end{aligned}$$

As this result, and noting that $\cos(\pi y_k) - 1 = -2 \sin^2 \frac{k\pi h}{2} \neq 0 (k = 1, 2, \dots, N-1)$, one we can obtain

$$\begin{aligned} \sum_{j=1}^{N-1} \varphi_j \sin(j\pi y_k) &= -\frac{1}{2 \sin^2 \frac{k\pi h}{2}} \sum_{j=2}^{N-2} \left(\frac{\varphi_{j-1} + \varphi_{j+1}}{2} - \varphi_j \right) \sin(j\pi y_k) \\ &\quad + \frac{\cos \frac{k\pi h}{2}}{\sin \frac{k\pi h}{2}} \left(-\left(\frac{1}{2} \varphi_2 - \varphi_1 \right) + (-1)^k \left(\frac{1}{2} \varphi_{N-2} - \varphi_{N-1} \right) \right). \end{aligned} \quad (4.14)$$

Using

$$\sin x \geq \frac{2}{\pi} x, \quad \forall x \in (0, \frac{\pi}{2}), \quad (4.15)$$

we get

$$\sin \frac{k\pi h}{2} \geq kh > 0. \quad (4.16)$$

From (4.14) and (4.16), it follows

$$\begin{aligned} \left| \sum_{j=1}^{N-1} \varphi_j \sin(j\pi y_k) \right| &\lesssim \frac{h^{-1}}{k \sin \frac{k\pi h}{2}} \sum_{j=2}^{N-2} \left| \frac{\varphi_{j-1} + \varphi_{j+1}}{2} - \varphi_j \right| \\ &\quad + \frac{1}{\sin \frac{k\pi h}{2}} (|\varphi_2| + |\varphi_1| + |\varphi_{N-2}| + |\varphi_{N-1}|) \\ &\lesssim \frac{h^{-1}}{k \sin \frac{k\pi h}{2}} \sum_{j=2}^{N-2} \left| \frac{\varphi_{j-1} + \varphi_{j+1}}{2} - \varphi_j \right| + \frac{\|\varphi\|_{\infty}}{\sin \frac{k\pi h}{2}}. \end{aligned} \quad (4.17)$$

Noting that $\varphi(y) \in C^2([0, 1])$ and by the Taylor series expansions of $\varphi_{j\pm 1}$ at $y = y_j$, we have

$$\left| \frac{\varphi_{j-1} + \varphi_{j+1}}{2} - \varphi_j \right| \lesssim h^2 \|\varphi''\|_\infty, \quad j = 2, 3, \dots, N-2. \quad (4.18)$$

Using (4.17), (4.18) and paying attention to $N = h^{-1}$, (4.13) holds. \square

Set the integer $K > 3$, $\beta \in \mathbb{R}$, a $K \times K$ square matrix is defined by

$$T = \begin{pmatrix} \beta & -1 & & \\ -1 & \beta & -1 & \\ & \ddots & \ddots & \ddots \\ & & -1 & \beta \end{pmatrix}. \quad (4.19)$$

For any given real number λ , let

$$Z_j(\lambda) = \lambda^j - \lambda^{-j}, \quad j = 0, 1, 2, \dots. \quad (4.20)$$

If $|\lambda| > 1$, paying attention to $Z_0(\lambda) = 0$ and $|Z_j(\lambda)| = |\lambda|^j - |\lambda|^{-j}$, we have

$$\frac{|Z_j(\lambda)|}{|Z_i(\lambda)|} < 1, \quad 1 \leq j < i. \quad (4.21)$$

The following lemma presents the estimation of the elements of T^{-1} , where T^{-1} is the inverse of T .

Lemma 4.2. *If $|\beta| > 2$, T is defined by (4.19) and $T^{-1} := (\tilde{t}_{i,j})_{K \times K}$, then*

$$|\tilde{t}_{i,i}| = \max_j |\tilde{t}_{i,j}|, \quad i = 1, 2, \dots, K, \quad (4.22)$$

and

$$\tilde{t}_{K,K} < 1. \quad (4.23)$$

Proof. Since $TT^{-1} = I$, for any given $i = 1, 2, \dots, K$, it follows

$$\begin{cases} \beta \tilde{t}_{i,1} - \tilde{t}_{i,2} = \delta_{i,1}, \\ -\tilde{t}_{i,j-1} + \beta \tilde{t}_{i,j} - \tilde{t}_{i,j+1} = \delta_{i,j}, \quad j = 2, 3, \dots, K-1, \\ -\tilde{t}_{i,K-1} + \beta \tilde{t}_{i,K} = \delta_{i,K}. \end{cases} \quad (4.24)$$

Noticing that the characteristic equation of (4.24) is

$$\lambda^2 - \beta\lambda + 1 = 0. \quad (4.25)$$

Using $|\beta| > 2$, we get the characteristic root of equation (4.25) with its absolute value greater than 1 as follows

$$\lambda_\beta = \text{sgn}(\beta) \left(\frac{|\beta|}{2} + \sqrt{\left(\frac{\beta}{2}\right)^2 - 1} \right), \quad (4.26)$$

where $\text{sgn}(\cdot)$ is the sign function.

By (4.26), we can easily confirm that the component solution of difference equations (4.24) is

$$\tilde{t}_{i,j} = \begin{cases} \frac{Z_{K-i+1}(\lambda_\beta)Z_j(\lambda_\beta)}{Z_{K+1}(\lambda_\beta)Z_1(\lambda_\beta)}, & \text{if } j = 1, 2, \dots, i, \\ \frac{Z_i(\lambda_\beta)Z_{K-j+1}(\lambda_\beta)}{Z_{K+1}(\lambda_\beta)Z_1(\lambda_\beta)}, & \text{if } j = i+1, i+2, \dots, K. \end{cases} \quad (4.27)$$

Hence, using (4.27), we get

$$\frac{|\tilde{t}_{i,j}|}{|\tilde{t}_{i,i}|} = \begin{cases} \frac{|Z_j(\lambda_\beta)|}{|Z_i(\lambda_\beta)|}, & \text{if } j = 1, 2, \dots, i, \\ \frac{|Z_{K-j+1}(\lambda_\beta)|}{|Z_{K-i+1}(\lambda_\beta)|}, & \text{if } j = i+1, i+2, \dots, K. \end{cases} \quad (4.28)$$

Noticing that $|\lambda_\beta| > 1$, from (4.21) and (4.28), we have (4.22).

By using (4.20), (4.27) and (4.26), we have

$$\beta \tilde{t}_{i,i} = \beta \operatorname{sgn}(\lambda_\beta) \frac{|Z_{K-i+1}(\lambda_\beta)||Z_i(\lambda_\beta)|}{|Z_{K+1}(\lambda_\beta)||Z_1(\lambda_\beta)|} > 0. \quad (4.29)$$

From (4.24), (4.22), (4.29) and $|\beta| > 2$, then we have

$$1 = -\tilde{t}_{K,K-1} + \beta \tilde{t}_{K,K} \geq -|\tilde{t}_{K,K}| + \beta \tilde{t}_{K,K} \geq (|\beta| - 1)|\tilde{t}_{K,K}| > |\tilde{t}_{K,K}|.$$

That is to say (4.23) holds. □

For any given vector $\vec{\alpha} = (\alpha_1, \dots, \alpha_K)^T \in \mathbb{R}^K$, we consider

$$T\vec{s} = \vec{\alpha}, \quad (4.30)$$

where T is defined by (4.19) and $\vec{s} = (s_1, \dots, s_K)^T$.

Corollary 4.1. *Under the assumptions as Lemma 4.2, we have the last component of the solution vector of (4.30) satisfies*

$$|s_K| < K \|\vec{\alpha}\|_\infty. \quad (4.31)$$

Proof. Using $\vec{s} = T^{-1}\vec{\alpha}$, (4.22) and (4.23), we have

$$|s_K| = \left| \sum_{j=1}^K \tilde{t}_{K,j} \alpha_j \right| \leq \|\vec{\alpha}\|_\infty \sum_{j=1}^K |\tilde{t}_{K,j}| \leq \|\vec{\alpha}\|_\infty K |\tilde{t}_{K,K}| < \|\vec{\alpha}\|_\infty K,$$

which completes the proof of (4.31). □

Using the above lemmas and corollary, we can get the estimation of $\|\vec{e}_k^{(1)}\|_\infty (k = 1, 2, \dots, N-1)$.

Theorem 4.1. *If $u \in W$, then*

$$\|\vec{e}_k^{(1)}\|_\infty \lesssim \frac{h^{\frac{3}{2}}}{k}, \quad k = 1, 2, \dots, N-1, \quad (4.32)$$

and

$$|\vec{e}_{M-1,k}^{(1)}| \lesssim \frac{h^{\frac{7}{2}}}{\sin \frac{k\pi h}{2} |\cos(k\pi h)|}, \quad |\vec{e}_{M+2,k}^{(1)}| \lesssim \frac{h^{\frac{7}{2}}}{\sin \frac{k\pi h}{2} |\cos(k\pi h)|}. \quad (4.33)$$

Proof. Noticing that the inequation of (4.32) holds if and only if

$$\max_{0 \leq i \leq M} |\bar{e}_{i,k}^{(1)}| \lesssim \frac{h^{\frac{3}{2}}}{k} \text{ and } \max_{M+1 \leq i \leq N} |\bar{e}_{i,k}^{(1)}| \lesssim \frac{h^{\frac{3}{2}}}{k}. \quad (4.34)$$

Now we turn to prove the first inequation of (4.33) and (4.34), respectively, the second inequation of (4.33) and (4.34) can be proved similarly.

For any given $k = 1, 2, \dots, N-1$ and using (4.11), one can obtain that $\{\bar{e}_{i,k}^{(1)}\}_{i=0}^M$ satisfies the following difference equations

$$\begin{cases} \bar{L}_h \bar{e}_{i,k}^{(1)} = \bar{R}_{i,k}, & i = 1, 2, \dots, M-1, \\ \bar{e}_{0,k}^{(1)} = \bar{e}_{M,k}^{(1)} = 0, \end{cases} \quad (4.35)$$

where $\bar{L}_h \bar{e}_{i,k}^{(1)} := -\kappa^- \bar{e}_{i-1,k}^{(1)} \cos(k\pi h) + (2\kappa^- + h^2) \bar{e}_{i,k}^{(1)} - \kappa^- \bar{e}_{i+1,k}^{(1)} \cos(k\pi h)$.

Comparing (4.35) with (4.30), and using Corollary 4.1, where $s_i := \kappa^- \bar{e}_{i,k}^{(1)} \cos(k\pi h)$, $|\beta| := |\frac{2\kappa^- + h^2}{\kappa^- \cos(k\pi h)}| > 2$, $\alpha_i := \bar{R}_{i,k}$ and $K := M-1$, we get

$$|\kappa^- \bar{e}_{M-1,k} \cos(k\pi h)| = |s_K| < K \|\vec{\alpha}\|_\infty < M \max_{1 \leq i \leq M-1} |\bar{R}_{i,k}|.$$

Hence, by using $\kappa^- \geq 1$ (see (2.8)), we have

$$|\bar{e}_{M-1,k}| < \frac{M \max_{1 \leq i \leq M-1} |\bar{R}_{i,k}|}{\kappa^- |\cos(k\pi h)|}. \quad (4.36)$$

Noticing that $u \in W$, from (4.7) and (3.2), we have

$$|\bar{R}_{i,k}| \lesssim h^{\frac{9}{2}} \left| \sum_{j=1}^{N-1} C^{(4)}(x_i, y_j) \sin(j\pi y_k) \right| + h^{\frac{11}{2}}, \quad i = 1, 2, \dots, M-1. \quad (4.37)$$

Using (3.3) and Lemma 4.1, we get

$$\left| \sum_{j=1}^{N-1} C^{(4)}(x_i, y_j) \sin(j\pi y_k) \right| \lesssim \frac{1}{\sin \frac{k\pi h}{2}}, \quad i = 1, 2, \dots, M-1.$$

Substituting this into (4.37), we can obtain

$$|\bar{R}_{i,k}| \lesssim \frac{h^{\frac{9}{2}}}{\sin \frac{k\pi h}{2}}, \quad i = 1, 2, \dots, M-1. \quad (4.38)$$

Using (4.36) and (4.38), we have the first inequation of (4.33) is established.
Let

$$E_j = (1 + x_j) h^{-2} \max_{1 \leq i \leq M-1} |\bar{R}_{i,k}|, \quad j = 0, 1, \dots, M. \quad (4.39)$$

Noticing that $\bar{L}_h(1 + x_j) > h^2$, together with (4.39) and (4.35), we obtain

$$\bar{L}_h E_j > \max_{1 \leq i \leq M-1} |\bar{R}_{i,k}| \geq |\bar{L}_h \bar{e}_{j,k}^{(1)}|, \quad j = 1, 2, \dots, M-1. \quad (4.40)$$

Thus, using (4.40), and paying attention to $E_0 > |\bar{e}_{0,k}^{(1)}|$, $E_M > |\bar{e}_{M,k}^{(1)}|$, by comparison theorem, we have

$$|\bar{e}_{j,k}^{(1)}| < E_j \lesssim h^{-2} \max_{1 \leq i \leq M-1} |\bar{R}_{i,k}|, \quad j = 1, 2, \dots, M-1. \quad (4.41)$$

Hence, from (4.41), (4.38) and (4.16), the first inequation of (4.34) holds.
The proof is completed. \square

Then, we will estimate $\|\bar{e}_k^{(2)}\|_\infty (k = 1, 2, \dots, N-1)$.

Using (4.6), (4.7) and (4.8), similar to the derivation process of (4.11), we can obtain the 1D difference equations that $\bar{e}_k^{(2)} (k = 1, 2, \dots, N-1)$ satisfies as follows

$$\begin{cases} -\kappa^- \bar{e}_{i-1,k}^{(2)} \cos(k\pi h) + (2\kappa^- + h^2) \bar{e}_{i,k}^{(2)} - \kappa^- \bar{e}_{i+1,k}^{(2)} \cos(k\pi h) = 0, & i = 1, 2, \dots, M-1, \\ -\bar{e}_{i-1,k}^{(2)} \cos(k\pi h) + (2 + h^2) \bar{e}_{i,k}^{(2)} - \bar{e}_{i+1,k}^{(2)} \cos(k\pi h) = 0, & i = M+2, M+3, \dots, N-1, \\ -\kappa^- \bar{e}_{M-1,k}^{(2)} \cos(k\pi h) + (\kappa^- + \kappa^* + h^2) \bar{e}_{M,k}^{(2)} - \kappa^* \bar{e}_{M+1,k}^{(2)} \cos(k\pi h) = \bar{\gamma}_{M,k}, \\ -\kappa^* \bar{e}_{M,k}^{(2)} \cos(k\pi h) + (\kappa^* + 1 + h^2) \bar{e}_{M+1,k}^{(2)} - \bar{e}_{M+2,k}^{(2)} \cos(k\pi h) = \bar{\gamma}_{M+1,k}, \\ \bar{e}_{0,k}^{(2)} = \bar{e}_{N,k}^{(2)} = 0, \end{cases} \quad (4.42)$$

where

$$\bar{\gamma}_{M,k} = \bar{R}_{M,k} + \kappa^- \bar{e}_{M-1,k}^{(1)} \cos(k\pi h), \quad \bar{\gamma}_{M+1,k} = \bar{R}_{M+1,k} + \bar{e}_{M+2,k}^{(1)} \cos(k\pi h). \quad (4.43)$$

Lemma 4.3. *If function Z_j is defined by (4.20), then, for any given $k = 1, 2, \dots, N-1$, $\bar{e}_k^{(2)}$ satisfies*

$$\bar{e}_{M,k}^{(2)} = \frac{\bar{\gamma}_{M,k} \delta_k(1) + \bar{\gamma}_{M+1,k} \kappa^*}{(\delta_k(\kappa^-) \delta_k(1) - (\kappa^*)^2) \cos(k\pi h)}, \quad \bar{e}_{M+1,k}^{(2)} = \frac{\bar{\gamma}_{M+1,k} + \kappa^* \cos(k\pi h) \bar{e}_{M,k}^{(2)}}{\delta_k(1) \cos(k\pi h)}, \quad (4.44)$$

and

$$\bar{e}_{i,k}^{(2)} = \begin{cases} \frac{Z_i(\lambda_k(\kappa^-))}{Z_M(\lambda_k(\kappa^-))} \bar{e}_{M,k}^{(2)}, & \text{if } i = 0, 1, \dots, M-1, \\ \frac{Z_{N-i}(\lambda_k(1))}{Z_M(\lambda_k(1))} \bar{e}_{M+1,k}^{(2)}, & \text{if } i = M+2, M+3, \dots, N, \end{cases} \quad (4.45)$$

where $(\omega = \kappa^-, 1)$

$$\lambda_k(\omega) = \text{sgn}(\cos(k\pi h)) \left(\frac{1}{|\cos(k\pi h)|} \left(1 + \frac{h^2}{2\omega} \right) + \sqrt{\frac{1}{\cos^2(k\pi h)} \left(1 + \frac{h^2}{2\omega} \right)^2 - 1} \right), \quad (4.46)$$

$$\delta_k(\omega) = \Theta_k(\omega) + \frac{\kappa^* + h^2}{\cos(k\pi h)}, \quad \Theta_k(\omega) = \omega \left(\frac{1}{\cos(k\pi h)} - \frac{Z_{M-1}(\lambda_k(\omega))}{Z_M(\lambda_k(\omega))} \right). \quad (4.47)$$

Proof. Using the former $2M-2$ equations of (4.42), and $\bar{e}_{0,k}^{(2)} = \bar{e}_{N,k}^{(2)} = 0$, we have

$$\begin{cases} (2\kappa^- + h^2) \bar{e}_{1,k}^{(2)} - \kappa^- \bar{e}_{2,k}^{(2)} \cos(k\pi h) = 0, \\ -\kappa^- \bar{e}_{i-1,k}^{(2)} \cos(k\pi h) + (2\kappa^- + h^2) \bar{e}_{i,k}^{(2)} - \kappa^- \bar{e}_{i+1,k}^{(2)} \cos(k\pi h) = 0, & i = 2, 3, \dots, M-2, \\ -\kappa^- \bar{e}_{M-2,k}^{(2)} \cos(k\pi h) + (2\kappa^- + h^2) \bar{e}_{M-1,k}^{(2)} = \kappa^- \bar{e}_{M,k}^{(2)} \cos(k\pi h), \end{cases} \quad (4.48)$$

and

$$\begin{cases} (2 + h^2) \bar{e}_{M+2,k}^{(2)} - \bar{e}_{M+3,k}^{(2)} \cos(k\pi h) = \bar{e}_{M+1,k}^{(2)} \cos(k\pi h), \\ -\bar{e}_{i-1,k}^{(2)} \cos(k\pi h) + (2 + h^2) \bar{e}_{i,k}^{(2)} - \bar{e}_{i+1,k}^{(2)} \cos(k\pi h) = 0, & i = M+3, M+4, \dots, N-2, \\ -\bar{e}_{N-2,k}^{(2)} \cos(k\pi h) + (2 + h^2) \bar{e}_{N-1,k}^{(2)} = 0. \end{cases} \quad (4.49)$$

By using (4.48), similar to the derivation process of (4.27) ($\beta := \frac{2\kappa^- + h^2}{\kappa^- \cos(k\pi h)}$, $\lambda_\beta := \lambda_k(\kappa^-)$), we have

$$\bar{e}_{i,k}^{(2)} = \frac{Z_i(\lambda_k(\kappa^-))}{Z_M(\lambda_k(\kappa^-))} \bar{e}_{M,k}^{(2)}, \quad i = 0, 1, \dots, M-1.$$

Similarly, by (4.49), we have

$$\bar{e}_{i,k}^{(2)} = \frac{Z_{N-i}(\lambda_k(1))}{Z_M(\lambda_k(1))} \bar{e}_{M+1,k}^{(2)}, \quad i = M+2, M+3, \dots, N,$$

which completes the proof of (4.45).

Then using the $2M-1$ -th and $2M$ -th equation of (4.45), one can obtain

$$\begin{cases} \delta_k(\kappa^-) \cos(k\pi h) \bar{e}_{M,k}^{(2)} - \kappa^* \cos(k\pi h) \bar{e}_{M+1,k}^{(2)} = \bar{\gamma}_{M,k}, \\ -\kappa^* \cos(k\pi h) \bar{e}_{M,k}^{(2)} + \delta_k(1) \cos(k\pi h) \bar{e}_{M+1,k}^{(2)} = \bar{\gamma}_{M+1,k}, \end{cases}$$

where $\delta_k(\omega)$ is defined by (4.47).

Solving the above equations, we deduce that $\bar{e}_{M,k}^{(2)}$ and $\bar{e}_{M+1,k}^{(2)}$ satisfy (4.44). \square

In order to estimate $\bar{e}_k^{(2)}$ by (4.44) and (4.45), we introduce the following two lemmas.

Lemma 4.4.

$$\delta_k(\kappa^-) \geq \delta_k(1) > \kappa^* + kh, \quad \text{if } k = 1, 2, \dots, M, \quad (4.50)$$

$$-\delta_k(\kappa^-) \geq -\delta_k(1) > \kappa^* + (N-k)h, \quad \text{if } k = M+1, M+2, \dots, N-1. \quad (4.51)$$

Proof. We only give the proof of (4.50), and (4.51) can be proved similarly.

First of all, we have to prove

$$\delta_k(\kappa^-) \geq \delta_k(1), \quad k = 1, 2, \dots, M. \quad (4.52)$$

In fact, from (4.47), we know that if we can prove

$$\Theta_k(\kappa^-) \geq \Theta_k(1), \quad k = 1, 2, \dots, M, \quad (4.53)$$

then (4.52) follows immediately.

Using (4.46) and

$$0 < \cos(k\pi h) < 1, \quad k = 1, 2, \dots, M, \quad (4.54)$$

we have

$$\frac{\omega}{\cos(k\pi h)} = \omega \lambda_k(\omega) - \frac{h^2}{2 \cos(k\pi h)} - \eta(\omega), \quad (4.55)$$

where $\eta(\omega) = \sqrt{(\frac{2\omega+h^2}{2 \cos(k\pi h)})^2 - \omega^2}$.

Hence, from (4.47) and (4.55), we have

$$\Theta_k(\omega) = \omega(\lambda_k(\omega) - \frac{Z_{M-1}(\lambda_k(\omega))}{Z_M(\lambda_k(\omega))}) - \frac{h^2}{2 \cos(k\pi h)} - \eta(\omega). \quad (4.56)$$

By (4.20) and (4.46), it follows that

$$w(\lambda_k(\omega) - \frac{Z_{M-1}(\lambda_k(\omega))}{Z_M(\lambda_k(\omega))}) = w \frac{\lambda_k(\omega) - (\lambda_k(\omega))^{-1}}{1 - (\lambda_k(\omega))^{-2M}} = \frac{2\eta(\omega)}{1 - (\lambda_k(\omega))^{-2M}}. \quad (4.57)$$

Thus, substituting (4.57) into (4.56), we have

$$\Theta_k(\omega) = (\frac{2}{1 - (\lambda_k(\omega))^{-2M}} - 1)\eta(\omega) - \frac{h^2}{2 \cos(k\pi h)}. \quad (4.58)$$

From (4.46) and $\kappa^- \geq 1$, one can obtain

$$\lambda_k(1) \geq \lambda_k(\kappa^-) > 1, \quad k = 1, 2, \dots, M. \quad (4.59)$$

Therefore, by using (4.58), (4.59), the characteristic that $\eta(\omega)$ is increasing monotone with $w > 0$ and $\kappa^- \geq 1$, we have (4.53) established.

Then, we want to illustrate

$$\delta_k(1) - \kappa^* > kh, \quad k = 1, 2, \dots, M. \quad (4.60)$$

Using (4.47) and (4.54), we get

$$\delta_k(1) - \kappa^* > \Theta_k(1). \quad (4.61)$$

From (4.58) and $\lambda_k(1) > 1$, we have

$$\Theta_k(1) > \eta(1) - \frac{h^2}{2 \cos(k\pi h)}. \quad (4.62)$$

For $k = 1, 2, \dots, M$ and using (4.15), it follows

$$\frac{1}{\cos^2(k\pi h)} > 1 + 4k^2h^2 \quad \text{and} \quad \frac{1}{\cos(k\pi h)} \leq h^{-1}, \quad k = 1, 2, \dots, M.$$

From this, one can obtain

$$\eta(1) \geq \sqrt{(1 + \frac{h^2}{2})^2(1 + 4k^2h^2) - 1} > 2kh \quad \text{and} \quad \frac{h^2}{2 \cos(k\pi h)} \leq \frac{h}{2}. \quad (4.63)$$

Combining (4.61), (4.62) and (4.63), we get (4.60).

In conclusion, using (4.52) and (4.60), we finish the proof of (4.50). \square

By using Lemma 4.4, it is obvious that

$$|\delta_k(\kappa^-)| \geq |\delta_k(1)| > \kappa^*. \quad (4.64)$$

Lemma 4.5. *For any given $k = 1, 2, \dots, N-1$, if $u \in W$, then*

$$|\bar{\gamma}_{M+1,k}| \lesssim \frac{h^{\frac{3}{2}}}{\sin \frac{k\pi h}{2}}, \quad (4.65)$$

$$|\bar{\gamma}_{M,k}\delta_k(1) + \bar{\gamma}_{M+1,k}\kappa^*| \lesssim \frac{h^{\frac{3}{2}}}{\sin \frac{k\pi h}{2}} (|\delta_k(1) - \kappa^*| + h|\delta_k(1)|). \quad (4.66)$$

Further, if the Assumption **I** is established, then

$$|\bar{\gamma}_{M+1,k}| \lesssim \frac{h^{\frac{5}{2}}}{\sin \frac{k\pi h}{2}}, \quad (4.67)$$

$$|\bar{\gamma}_{M,k}\delta_k(1) + \bar{\gamma}_{M+1,k}\kappa^*| \lesssim \frac{h^{\frac{5}{2}}}{\sin \frac{k\pi h}{2}}(|\delta_k(1) - \kappa^*| + h|\delta_k(1)|). \quad (4.68)$$

Proof. Using (4.7), (3.4), (3.5) and Lemma 4.1, similar to the proof of (4.38) shows that

$$|\bar{R}_{M+1,k}| \lesssim \frac{h^{\frac{3}{2}}}{\sin \frac{k\pi h}{2}}, \quad k = 1, 2, \dots, N-1. \quad (4.69)$$

Similarly, if the Assumption **I** is established, by using (4.7), (3.14), (3.15) and Lemma 4.1, we have

$$|\bar{R}_{M+1,k}| \lesssim \frac{h^{\frac{5}{2}}}{\sin \frac{k\pi h}{2}}, \quad k = 1, 2, \dots, N-1. \quad (4.70)$$

Therefore, from (4.43), (4.33) and (4.69), we get (4.65). Similarly, using (4.70), we get (4.67).

The rest is to show that (4.66) and (4.68) are established.

Using (4.43) and (4.7), it follows

$$\begin{aligned} & |\bar{\gamma}_{M,k}\delta_k(1) + \bar{\gamma}_{M+1,k}\kappa^*| \\ & \leq |\bar{R}_{M,k}\delta_k(1) + \bar{R}_{M+1,k}\kappa^*| + \kappa^- |\bar{e}_{M-1,k}^{(1)} \cos(k\pi h)| |\delta_k(1)| + |\bar{e}_{M+2,k}^{(1)} \cos(k\pi h)| \kappa^* \\ & = \beta_1 + \kappa^- |\bar{e}_{M-1,k}^{(1)} \cos(k\pi h)| |\delta_k(1)| + |\bar{e}_{M+2,k}^{(1)} \cos(k\pi h)| \kappa^*, \end{aligned} \quad (4.71)$$

where

$$\beta_1 = \sqrt{2h} \left| \sum_{j=1}^{N-1} (R_{M,j}\delta_k(1) + R_{M+1,j}\kappa^*) \sin(j\pi y_k) \right|. \quad (4.72)$$

By (4.33) and Lemma 4.4, we have

$$\kappa^- |\bar{e}_{M-1,k}^{(1)} \cos(k\pi h)| |\delta_k(1)| + |\bar{e}_{M+2,k}^{(1)} \cos(k\pi h)| \kappa^* \lesssim |\delta_k(1)| \frac{h^{\frac{7}{2}}}{\sin \frac{k\pi h}{2}}. \quad (4.73)$$

Hence, substituting (4.73) into (4.71), we can obtain

$$|\bar{\gamma}_{M,k}\delta_k(1) + \bar{\gamma}_{M+1,k}\kappa^*| \lesssim \beta_1 + |\delta_k(1)| \frac{h^{\frac{7}{2}}}{\sin \frac{k\pi h}{2}}. \quad (4.74)$$

Using (4.72), (3.4) and (3.5), and noticing that (4.64), we can obtain

$$\begin{aligned}
\beta_1 &\lesssim h^{\frac{3}{2}} |\delta_k(1) - \kappa^*| \left| \sum_{j=1}^{N-1} C_M^{(1)}(y_j) \sin(j\pi y_k) \right| \\
&\quad + h^{\frac{5}{2}} \left| \sum_{j=1}^{N-1} (C_M^{(2)}(y_j) \delta_k(1) + C_{M+1}^{(2)}(y_j) \kappa^*) \sin(j\pi y_k) \right| + h^{\frac{5}{2}} |\delta_k(1)| \\
&\leq h^{\frac{3}{2}} |\delta_k(1) - \kappa^*| \left| \sum_{j=1}^{N-1} C_M^{(1)}(y_j) \sin(j\pi y_k) \right| \\
&\quad + h^{\frac{5}{2}} |\delta_k(1)| \left(\left| \sum_{j=1}^{N-1} C_M^{(2)}(y_j) \sin(j\pi y_k) \right| + \left| \sum_{j=1}^{N-1} C_{M+1}^{(2)}(y_j) \sin(j\pi y_k) \right| + h^{\frac{5}{2}} |\delta_k(1)| \right). \quad (4.75)
\end{aligned}$$

Further, if Assumption **I** holds, by similar derivation of (4.75), using (4.72), (3.14), (3.15), and noticing that (4.64) holds, we have

$$\begin{aligned}
\beta_1 &\lesssim h^{\frac{5}{2}} |\delta_k(1) - \kappa^*| \left| \sum_{j=1}^{N-1} \tilde{C}_M^{(2)}(y_j) \sin(j\pi y_k) \right| \\
&\quad + h^{\frac{7}{2}} |\delta_k(1)| \left(\left| \sum_{j=1}^{N-1} C_M^{(3)}(y_j) \sin(j\pi y_k) \right| + \left| \sum_{j=1}^{N-1} C_{M+1}^{(3)}(y_j) \sin(j\pi y_k) \right| + h^{\frac{7}{2}} |\delta_k(1)| \right). \quad (4.76)
\end{aligned}$$

Therefore, using (4.75), (3.5), (3.6), (3.7) and Lemma 4.1 we have

$$\beta_1 \lesssim |\delta_k(1) - \kappa^*| \frac{h^{\frac{3}{2}}}{\sin \frac{k\pi h}{2}} + |\delta_k(1)| \frac{h^{\frac{5}{2}}}{\sin \frac{k\pi h}{2}}. \quad (4.77)$$

Similarly, if Assumption **I** is established, from (4.76), (3.15), (3.8), (3.9) and Lemma 4.1, we have

$$\beta_1 \lesssim |\delta_k(1) - \kappa^*| \frac{h^{\frac{5}{2}}}{\sin \frac{k\pi h}{2}} + |\delta_k(1)| \frac{h^{\frac{7}{2}}}{\sin \frac{k\pi h}{2}}. \quad (4.78)$$

Consequently, substituting (4.77) and (4.78) into (4.74), respectively, we have (4.66) and (4.68) established. \square

Using the above lemmas, we can get the estimation of $\|\tilde{e}_k^{(2)}\|_\infty (k = 1, 2, \dots, N-1)$.

Theorem 4.2. *If $u \in W$, then*

$$\|\tilde{e}_k^{(2)}\|_\infty \lesssim h^{\frac{1}{2}} g(k), \quad k = 1, 2, \dots, N-1. \quad (4.79)$$

*Further, if the Assumption **I** holds, then*

$$\|\tilde{e}_k^{(2)}\|_\infty \lesssim h^{\frac{3}{2}} g(k), \quad k = 1, 2, \dots, N-1, \quad (4.80)$$

where

$$g(k) = g(N-k) = g_1(k), \quad k = 1, 2, \dots, M, \quad (4.81)$$

$$g_1(k) = \begin{cases} \frac{1}{k}, & \text{if } k = 1, 2, \dots, \lfloor \frac{N}{4} \rfloor, \\ \frac{1}{N-2k}, & \text{if } k = \lceil \frac{N}{4} \rceil, \lceil \frac{N}{4} \rceil + 1, \dots, M. \end{cases} \quad (4.82)$$

Proof. We only present the proof of (4.79), and (4.80) can be proved similarly.

Note that $|\lambda_k(\omega)| > 1$ ($\omega = \kappa^-, 1$), by (4.45) and (4.21), we have

$$|\bar{e}_{i,k}^{(2)}| \lesssim \max\{|\bar{e}_{M,k}^{(2)}|, |\bar{e}_{M+1,k}^{(2)}|\}, \quad i = 0, 1, \dots, M-1, M+2, M+3, \dots, N. \quad (4.83)$$

Then, we will present the estimation of $|\bar{e}_{M,k}^{(2)}|$ firstly.

Using (4.44) and (4.66), we have

$$|\bar{e}_{M,k}^{(2)}| \lesssim \frac{h^{\frac{3}{2}}}{\sin \frac{k\pi h}{2} |\cos(k\pi h)|} \left(\frac{|\delta_k(1) - \kappa^*|}{|\delta_k(\kappa^-)\delta_k(1) - (\kappa^*)^2|} + \frac{h|\delta_k(1)|}{|\delta_k(\kappa^-)\delta_k(1) - (\kappa^*)^2|} \right). \quad (4.84)$$

By the first inequation of (4.50) and (4.51), we have

$$\frac{1}{|\delta_k(\kappa^-)\delta_k(1) - (\kappa^*)^2|} \leq \frac{1}{|(\delta_k(1))^2 - (\kappa^*)^2|} = \frac{1}{|\delta_k(1) + \kappa^*||\delta_k(1) - \kappa^*|}.$$

From this, and using the second inequation of (4.50) and (4.51), we can obtain

$$\frac{|\delta_k(1) - \kappa^*|}{|\delta_k(\kappa^-)\delta_k(1) - (\kappa^*)^2|} \leq \frac{1}{|\delta_k(1) + \kappa^*|} \lesssim \begin{cases} 1, & \text{if } k = 1, 2, \dots, M, \\ \frac{1}{(N-k)h}, & \text{if } k = M+1, M+2, \dots, N-1, \end{cases} \quad (4.85)$$

and

$$\frac{|\delta_k(1)|}{|\delta_k(\kappa^-)\delta_k(1) - (\kappa^*)^2|} < \frac{1}{|\delta_k(1) - \kappa^*|} \lesssim \begin{cases} \frac{1}{kh}, & \text{if } k = 1, 2, \dots, M, \\ \frac{1}{(N-k)h}, & \text{if } k = M+1, M+2, \dots, N-1. \end{cases} \quad (4.86)$$

Substituting (4.85) and (4.86) into (4.84), we have

$$|\bar{e}_{M,k}^{(2)}| \lesssim h^{\frac{1}{2}} \tilde{g}(k), \quad (4.87)$$

where

$$\tilde{g}(k) = \begin{cases} \frac{h}{\sin \frac{k\pi h}{2} |\cos(k\pi h)|}, & \text{if } k = 1, 2, \dots, M, \\ \frac{1}{(N-k) \sin \frac{k\pi h}{2} |\cos(k\pi h)|}, & \text{if } k = M+1, M+2, \dots, N-1. \end{cases}$$

Using the basic characteristics of trigonometric function, it is easy to verify

$$\tilde{g}(k) \lesssim g(k), \quad (4.88)$$

where $g(k)$ is defined by (4.81).

Hence, substituting (4.88) into (4.87), we have

$$|\bar{e}_{M,k}^{(2)}| \lesssim h^{\frac{1}{2}} g(k). \quad (4.89)$$

Secondly, by (4.44), Lemma 4.4, (4.65), $\frac{h}{\sin \frac{k\pi h}{2} |\cos(k\pi h)|} \lesssim g(k)$ and (4.89), we have

$$|\bar{e}_{M+1,k}^{(2)}| \leq \frac{|\gamma_{M+1,k}| + \kappa^* |\cos(k\pi h)| |\bar{e}_{M,k}^{(2)}|}{|\delta_k(1)| |\cos(k\pi h)|} \lesssim \frac{1}{\sin \frac{k\pi h}{2} |\cos(k\pi h)|} h^{\frac{3}{2}} + |\bar{e}_{M,k}^{(2)}| \lesssim h^{\frac{1}{2}} g(k). \quad (4.90)$$

In conclusion, combining (4.83), (4.89) and (4.90), we finish the proof of (4.79). \square

The following theorem gives the estimation of $\|\vec{e}^{(l)}\|_\infty$ ($l = 1, 2$).

Theorem 4.3. *If $u \in W$, then*

$$\|\vec{e}^{(1)}\|_\infty \lesssim h^2 |\ln h|, \quad \|\vec{e}^{(2)}\|_\infty \lesssim h |\ln h|. \quad (4.91)$$

Further, if the Assumption I holds, then

$$\|\vec{e}^{(2)}\|_\infty \lesssim h^2 |\ln h|. \quad (4.92)$$

Proof. From (4.8), we have

$$|e_{i,j}^{(l)}| \leq \sqrt{2h} \sum_{k=1}^{N-1} |\bar{e}_{i,k}^{(l)}|, \quad l = 1, 2. \quad (4.93)$$

Using (4.81), we have

$$\sum_{k=1}^{N-1} g(k) = 2 \left(\sum_{k=1}^{\lfloor \frac{N}{4} \rfloor} \frac{1}{k} + \sum_{k=\lceil \frac{N}{4} \rceil}^M \frac{1}{N-2k} \right) \lesssim \sum_{k=1}^{N-1} \frac{1}{k} \lesssim |\ln h|. \quad (4.94)$$

Therefore, substituting (4.32), (4.79) and (4.80) into (4.93), respectively, and using (4.94), we have (4.91) and (4.92) established. \square

Using Theorem 4.3 and (4.5), we can obtain the main result of this paper.

Theorem 4.4. *If $u \in W$, then*

$$\|\vec{e}\|_\infty \lesssim h |\ln h|. \quad (4.95)$$

Further, if the Assumption I holds, then

$$\|\vec{e}\|_\infty \lesssim h^2 |\ln h|. \quad (4.96)$$

Remark 4.1. *Although the above theoretical estimations just prove that the Assumption I is the sufficient condition for scheme (2.10) to reach the optimal asymptotic order, the numerical experiments (see Table 1 and Table 2 in the following section) show that the assumptions also seems to be necessary.*

5. Numerical experiments

In this section, some typical numerical experiments are carried out to experimentally study the accuracy of the five-point MACH-like scheme (2.10).

Example 5.1. *Consider the simplified model of (2.1) and (2.2), where the diffusion coefficient $\kappa^- = 10^4$, the exact solution $u(x, y) = \sin(\pi x) \sin(\pi y) ((x - \frac{1}{2})/\kappa + 1)$ does not satisfy (3.13).*

The experiment results for solving Example 5.1 are given in Table 1, where $\tilde{\kappa}_1 = \frac{\kappa^- + 1}{2}$ and $\tilde{\kappa}_2 = \frac{2\kappa^-}{\kappa^- + 1}$. It can be observed that if the exact solution does not satisfy the assumption condition (3.13), no matter whether the diffusion coefficients in the multiple material cells use harmonic averaging or not, this scheme can not reach the asymptotic optimal error estimate $O(h^2 |\ln h|)$ in the maximum norm.

Table 1: Results for $u(x, y) = \sin(\pi x) \sin(\pi y)((x - \frac{1}{2})/\kappa + 1)$ and $\kappa^- = 10^4$.

| N | $\kappa^* = \tilde{\kappa}_1$ | | $\kappa^* = \tilde{\kappa}_2$ | |
|-----|-------------------------------|-------|-------------------------------|-------|
| | $\ \tilde{e}\ _\infty$ | ratio | $\ \tilde{e}\ _\infty$ | ratio |
| 33 | 1.65E-02 | — | 4.64E-02 | — |
| 67 | 7.81E-03 | 2.12 | 2.21E-02 | 2.09 |
| 135 | 3.79E-03 | 2.06 | 1.08E-02 | 2.05 |
| 271 | 1.87E-03 | 2.03 | 5.35E-03 | 2.02 |

Table 2: Results for $u(x, y) = \frac{1}{\kappa} \sin(\pi x) \sin(\pi y)(x - \frac{1}{2})(y - 1)y(1 + x^2 + y^2)$ and $\kappa^- = 10^6$.

| N | $\kappa^* = \tilde{\kappa}_1$ | | $\kappa^* = \tilde{\kappa}_2$ | |
|-----|-------------------------------|-------|-------------------------------|-------|
| | $\ \tilde{e}\ _\infty$ | ratio | $\ \tilde{e}\ _\infty$ | ratio |
| 33 | 5.80E-03 | — | 4.70E-04 | — |
| 67 | 2.85E-03 | 2.04 | 1.14E-04 | 4.12 |
| 135 | 1.41E-03 | 2.02 | 2.81E-05 | 4.06 |
| 271 | 7.02E-04 | 2.01 | 6.98E-06 | 4.03 |

Example 5.2. Consider the simplified model of (2.1) and (2.2), where the diffusion coefficient $\kappa^- = 10^6$, the exact solution $u(x, y) = \frac{1}{\kappa} \sin(\pi x) \sin(\pi y)(x - \frac{1}{2})(y - 1)y(1 + x^2 + y^2)$ satisfies (3.13).

The corresponding experiment results given in Table 2 show that only when the exact solution satisfies the assumption condition (3.13) and $\kappa^* = \tilde{\kappa}_2$, the scheme can reach the asymptotic optimal error estimate $O(h^2 |\ln h|)$, otherwise, the scheme can only reach $O(h |\ln h|)$.

All the results above-mentioned verify the correctness of the error theories.

In addition, a series of numerical experiments have carried out for the more general nine-point scheme (2.6). The experimental results are similar to those of the five-point scheme, but it is omitted on account of space limitation.

6. Conclusions

In this paper, a kind of vertex-centered MACH-like FVM for stationary diffusion problems with interface is constructed, and the estimates of the local truncation error and global error have been established, then the theoretical results are verified by numerical experiments. It's worth pointing out that, if the exact solution does not satisfy the assumption condition (3.13), the five-point MACH-like scheme with harmonic averaging can not reach the asymptotic optimal error estimate $O(h^2 |\ln h|)$ in the maximum norm. In the future, we hope to generalize this work to more complex diffusion problems and schemes, such as 3D diffusion problems and other schemes.

Acknowledgements

The authors would like to thank Dr. Cunyun Nie from Hunan Institute of Engineering for his helpful comments and suggestions. This work is supported by the National Natural Science Foundation of China (Grant Nos. 11571293, 11601462 and 61603322), Hunan Provincial Natural Science Foundation of China (Grant No. 2016JJ2129), and Open Foundation of Guangdong Provincial Engineering Technology Research Center for Data Science (Grant No. 2016KF03).

References

- [1] D. S. Kershaw, Differencing of the diffusion equation in lagrangian hydrodynamic codes, J. Comput. Phys. 39 (2) (1981) 375–395.

- [2] F. Hermeline, A finite volume method for the approximation of diffusion operators on distorted meshes, *J. Comput. Phys.* 160 (2000) 480–499.
- [3] J. Droniou, R. Eymard, A mixed finite volume scheme for anisotropic diffusion problems on any grid, *Numer. Math.* 105 (1) (2006) 35–71.
- [4] G. Yuan, Z. Sheng, Analysis of accuracy of a finite volume scheme for diffusion equations on distorted meshes, *J. Comput. Phys.* 224 (2007) 1170–1189.
- [5] R. A. Klausen, A. F. Stephansen, Convergence of multi-point flux approximations on general grids and media, *Int. J. Numer. Anal. Model.* 9 (3) (2012) 584–606.
- [6] J. Droniou, Finite volume schemes for diffusion equations: Introduction to and review of modern methods, *Math. Mod. Meth. Appl. S.* 24 (8) (2014) 1575–1619.
- [7] S. Kadioglu, R. R. Nourgaliev, V. A. Mousseau, A comparative study of the harmonic and arithmetic averaging of diffusion coefficients for non-linear heat conduction problems, Technical Report INL/EXT-08-13999, Idaho National Laboratory, Idaho Falls, Idaho 83415, 2008.
- [8] R. Ewing, Z. Li, T. Lin, Y. Lin, The immersed finite volume element methods for the elliptic interface problems, *Math. Comput. Simulat.* 50 (1) (1999) 63–76.
- [9] L. Zhu, Z. Zhang, Z. Li, An immersed finite volume element method for 2d pdes with discontinuous coefficients and non-homogeneous jump conditions, *Comput. Math. Appl.* 70 (2015) 89–103.
- [10] S. Shu, H. Y. Yu, Y. Q. Huang, C. Y. Nie, A symmetric finite volume element scheme on quadrilateral grids and superconvergence, *Int. J. Numer. Anal. Mod.* 3 (3) (2006) 348–360.
- [11] C. Y. Nie, S. Shu, H. Y. Yu, J. Wu, Superconvergence and asymptotic expansions for bilinear finite volume element approximations, *Numer. Math. Theor. Meth. Appl.* 6 (2) (2013) 408–423.
- [12] L. Mu, R. Jari, A recovery-based error estimate for nonconforming finite volume methods of interface problems, *Appl. Math. Comput.* 220 (2013) 63–74.
- [13] M. Oevermann, C. Scharfenberg, R. Klein, A sharp interface finite volume method for elliptic equations on cartesian grids, *J. Comput. Phys.* 228 (2009) 5184–5206.
- [14] M. Oevermann, R. Klein, A cartesian grid finite volume method for elliptic equations with variable coefficients and embedded interfaces, *J. Comput. Phys.* 219 (2) (2006) 749–769.
- [15] R. Ewing, O. Iliev, R. Lazarov, A modified finite volume approximation of second-order elliptic equations with discontinuous coefficients, *SIAM J. Sci. Comput.* 23 (4) (2001) 1335–1351.
- [16] R. Luce, S. Perez, A finite volume scheme for an elliptic equation with heterogeneous coefficients. application to a homogenization problem, *Appl. Numer. Math.* 38 (4) (2001) 427–444.
- [17] T. Wang, Z. Zhang, A compact finite volume method and its extrapolation for elliptic equations with third boundary conditions, *Appl. Math. Comput.* 264 (2015) 258–271.
- [18] J. Lv, Y. Li, L2 error estimate of the finite volume element methods on quadrilateral meshes, *Adv. Comput. Math.* 33 (2010) 129–148.

- [19] J. Lv, Y. Li, Optimal biquadratic finite volume element methods on quadrilateral meshes, *SIAM J. Numer. Anal.* 50 (2012) 2379–2399.
- [20] C. Nie, S. Shu, H. Yu, Q. An, A high order composite scheme for the second order elliptic problem with nonlocal boundary and its fast algorithm, *Appl. Math. Comput.* 227 (2014) 212–221.
- [21] M. H. Frese, Mach2: A two-dimensional magnetohydrodynamic simulation code for complex experimental configurations, No. AMRC-R-874, Mission Research Corp., Albuquerque, NM, 1987.
- [22] U. Shumlak, T. W. Hussey, G. J. Marklin, R. E. Peterkin, Mach3: A three-dimensional mhd code, *Bull. Am. Phys. Soc.* 38 (10).
- [23] U. Shumlak, T. W. Hussey, R. E. Peterkin, Three-dimensional magnetic field enhancement in a liner implosion system, *IEEE Trans. Plas. Sci.* 23 (1) (1995) 83–88.
- [24] R. E. Peterkin, M. H. Frese, C. R. Sovinec, Transport of magnetic flux in an arbitrary coordinate ale code, *J. Comput. Phys.* 140 (1998) 148–171.
- [25] M. R. LaPointe, P. G. Mikellides, High-power electromagnetic thruster being developed, *Research & Technology* 2001, NASA/TM-2002-211333 (2002) 51–53.
- [26] P. G. Mikellides, Modeling and analysis of a megawatt-class magnetoplasmdynamic thruster, *J. Propul. Power* 20 (2) (2004) 204–210.
- [27] D. Ahern, Investigation of the gem mpd thruster using the mach2 magnetohydrodynamics code, University of Illinois at Urbana-Champaign, 2013.
- [28] D. Li, H. Shui, M. Tang, On the finite difference scheme of two-dimensional parabolic equation in a non-rectangular mesh, *J. Numer. Methods Comput. Appl.* 4 (1980) 217–224.
- [29] S. N. Barkas, An introduction to fast poisson solvers, *Philips Journal Research* 37 (5–6) (2005) 231–264.
- [30] M. V. Pham, F. Plourde, S. D. Kim, Strip decomposition parallelization of fast direct poisson solver on a 3d cartesian staggered grid, *Int. J. Comput. Sci. Eng.* 1 (3) (2007) 183–192.
- [31] M. X. Kan, G. H. Wang, H. L. Zhao, L. Xie, Two-dimensional magneto-hydrodynamic simulations of magnetically accelerated flyer plates, *High Power Laser and Particle Beams* 8 (2013) 052.